Measurement of the Performance of IP Packets over ATM Environment via Multiprotocol-over-ATM (MPOA)

Dipl.-Ing. Kai-Oliver Detken
Director *wwl network*, Bremen

WWL Internet AG,
Goebelstr. 46, D-28865 Lilienthal/Bremen, Germany
kai.detken@wwl.de,
Business: http://wwl.de
Private: http://kai.nord.de

Abstract. Multiprotocol-over-ATM (MPOA) is a service which supports layer-3 internetworking for hosts attached to ELANs (running a LAN Emulation Client-LEC), hosts attached to ATM networks, and hosts attached to legacy LANs. MPOA provides and delivers the functions of a router and takes advantage of the underlying ATM network as much as possible. That means, MPOA works as a virtual router with real QoS features of ATM. Furthermore, the integration of existing network with reduced protocol overhead and reduced equipment is a crucial point in terms of network performance cost efficiency. But MPOA has already some disadvantages, because MPOA is a very complex technology and the success of MPOA is doubtful. IP might be worth the complexity because it is so widely used, but it can be doubted if this holds true for other layer 3 protocols as well. Nevertheless, the MPOA model is a very promising technology which has also negative aspects like very complex software implementation (hard to handle) and QoS can only be supported if you use ATM network interface cards (NICs) directly without Ethernet. This paper will discuss the measurement results (latency, throughput) from the Internet Protocol (IP) in ATM and Ethernet environment via MPOA. Furthermore this paper will give an overview about available ATM components regarding MPOA and the further development of this technology.

1. Introduction

Witnessing this latest round of ATM switch tests at the WWL, it is difficult to compare ATM core switch products against those from the Gigabit-Ethernet camp, because of the fast, simple and efficient technology from Gigabit-Ethernet. However, ATM is not a dead issue; it's just not in favor anymore. The reason is, ATM is too hard to use. A big disappointment, however, was support for the Multiprotocol-over-ATM (MPOA) specification. Originally, the WWL wanted to measure several components from different manufactures and vendors. But out of more than eight invites sent out, only two vendors was ready to come into the lab and demonstrate an MPOA solution. These vendors were Cisco Systems and Cabeltron Systems (today: Entera-

S. Rao and K.I. Sletta (Eds.): INTERWORKING 2000, LNCS 1938, pp. 110-121, 2000.

sys). What made MPOA so important to companies and customers is that this was the great hope of the ATM Forum to integrate legacy network protocols with ATM's cell traffic structure. Using an MPOA server, network managers would in theory be able to use all the Layer 3 switching features common to Layer-3-Ethernet switches, plus a few other goodies.

For example, MPOA finally made it feasible to route ATM and other protocols into the same VLAN. If this seems like the MPOA server would become a traffic bottleneck in large networks, it would be wrong. MPOA also makes use of PNNI to provide wire-speed routing across the entire ATM segment simply by establishing direct client-to-client connections when congestion gets ugly. The WWL used for the tests different packet sizes, because of exact measuring of performance and latency. Fragmentation cause by to big packets, influence the measuring results dramatically. Typical IP packet sizes are 64 byte, 512 byte, and 1518 byte. Small packets are usual for client-server communication regarding databases. This load is normal for a network and decribe the daily operation.

On the other hand, packets with a size of 512 byte are typical for the smallest IP packet size. Additionally, 1518 byte fits exactly in the maximum size of an Ethernet frame. The packets have not to fragment in smaller packets. For automatically connections in LANE/MPOA scenarios, the LEC/MPOA client must communicate with the LES/MPOA server. Between this communication an additional measurement equipment (in our case: WinPharaoh from GN Nettest) was used. Additionally, the traffic was produce from an analyser (in our case: Smartbits 2000 (SMB-2000) from NetCom) at the edge of the network on the Ethernet site. Thereby we was able to measure the traffic and the performance end-to-end.

By the test scenario in Figure 1 the WWL was able to test the following parameters:

- Performance Ethernet-Ethernet
- Latency
- Frame loss

MPOA was used with its specific protocols like P-NNI, LANEv2, NHRP. For the test of the shortcut functionality the MPC devices got different values for the Shortcut Frame Count, especially at the Cabletron equipment. This parameter is the threshold for the number of frames per second and is responsible for the establishing of shortcuts. We recommend an value of 10 for a fast establishing of shortcuts. Normally the value has the number 65000.

This paper will show by the use of two different manufacturers that MPOA is working, fast, sufficient, and efficient. But also, the paper will show the weakness of the products or the MPOA approach. Additionally, we tested quality-of-service (QoS) features with NICs from Fore Systems and Olicom. Therefore, mainly two vendors Cisco and Cabletron was tested at the WWL for our customer and a evaluation of this technologies, which has an deep impact of the development of QoS and Multi-Protocol-Label-Switching (MPLS).

2. Comparison of LANE and MPOA

For the adaptation or integration of IP into ATM, there are different methods existing developed by the ATM-Forum (LANE and MPOA) and IETF (CLIP and Multiprotocol Label Switching). For this paper, LANE and MPOA have been tested compared with CLIP, because they are based on each other, they are in ATM products currently available and support legacy LANs such as Ethernet and Token Ring.

LAN Emulation (LANE) is after CLIP the second solution for IP-over-ATM and can be best characterized as a service developed by the ATM Forum that will enable existing LAN applications to be run over an ATM network. In order to do so, this service has to amalgamated the characteristics and behaviors of traditional Ethernet, Token Ring and FDDI networks. Moreover, it has to support a connectionless service as current LAN stations send data without establishing a connection before the operation takes place. Therefore, LANE must support broadcast and multicast traffic such as the kind of traffic allowed over shared media LANs. It has to allow the interconnection of traditional LANs with the amalgamated LAN and at the same time to maintain the MAC address identity associated with each individual device that is attached to a LAN. Finally, it has to protect the vast installed basis of existing LAN applications and enable them to work in the same way as before over an ATM network. This may be put into practise over the OSI layer 2, whereas the LANE technology can be understood as a bridging technology.

MPOA clients established VCCs with the MPOA server components in order to forward data packets or to request information. Using these components, the client is able to establish a more direct path. MPOA supports various kinds of routable protocols (IP, IPX, AppleTalk, etc.), integrates existing internetworking protocols (RFC-1577, RFC-1483, NHRP, MARS, RSVP) and the IETF and ATM Forum solutions (LANE, P-NNI) into a virtual router environment. Yet actually, MPOA supports only IP and based on LANE, RFC-1483, and NHRP. MPOA is a service with layer 3 internetworking support for hosts attached to ELANs, ATM networks and legacy LANs. Thus, the real premise behind MPOA is to provide and deliver the functionality of a router and to take as much advantage of the underlying ATM network as possible. MPOA works as a virtual router on the OSI layer 3. [1]

2.1 TCP/IP-over-ATM

IP is an important protocol used to achieve interoperability in a heterogeneous network. In contrast, the effectiveness of IP in high speed networks is not widely known. There is much doubt about the TCP performance, in particular, as the acknowledgement mechanisms, overhead size and parameter setting are considered to be obstacles. UDP is more appropriate for real-time data streams, but does not have any security mechanisms.

The TCP/IP protocol stacks were not designed for high speed performance networks in the first place. Several extensions of TCP protocols have been suggested in order to achieve higher performance over these networks and to improve the performance of connections with a high bandwidth delay. New bandwidth-intensive applications such as multimedia conferencing systems and characteristics of high speed networks have triggered research on advanced transport protocols. Therefore, the discovery of additional error sources is not surprising. Following, the reader will find factors identified by WWL that are responsible for inefficiencies of TCP protocols over ATM: [3]

- Send and receive socket buffer size
- Network: Maximum Transport Unit (MTU)
- Protocol: Maximum Segment Size (MSS)
- Transmitter: use of Nagle's algorithm
- Round Trip Time (RTT)
- Receiver: delayed acknowledgement mechanisms
- Transmitter: Silly Window Syndrome (SWS)
- Copy strategy at the socket interface
- Network congestion and lost notice

2.2 Scenarios for MPOA

The following ATM devices were used during the tests in order to measure the performance of MPOA and LANE:

1. Smart Switch 2200 (MPOA Client MPC): Firmware Release No.: 40726; 24 x 10/100 Mbps Ethernet Ports (not modular); 1 x FE-100TX(10/100 Mbps Ethernet Module); 1 x APIM-21R (ATM-Module)
2. Smart Switch 2000 (MPOA Server MPS): Firmware Release No.: 40771; 1 x FE-100TX(10/100 Mbps Ethernet Module); 1 x APIM-21R(ATM-Module); Device Hardware Revision: 00C; Device Firmware Revision: MPS_1.0.1; Device BOOTPROM Revision: 01.05.02
3. Smart Switch Router 2000 (MPOA Router): Router Device connected to MPS via IEEE 802.1Q; Firmware Release : ssr22a2; 2 x 8 10/100 Mbps Ethernet Ports (not modular); Software Version: 2.2.A.2; Boot Prom Version: prom-1.1.0.5; System Type: SSR 2000, Rev. 0; CPU Module Type: CPU-SSR2, Rev. 0; Processor: R5000, Rev 2.1, 159.99 MHz
4. Smart Switch 2500: ATM Switch with LECS, LES/BUS device; Firmware Release: 02.03(35); 3 x 4 OC-3 Ports (ZX-IOM-21-4); CPU Model: i960 CX; CPU Speed : 33 MHz
5. Catalyst 5506 (LEC ELAN 1, MPC-A): 100BaseTX Supervisor (Hardware 2.3; Software 4.5(3)); 10/100BaseTX Ethernet (Hardware 1.2; Software 4.5(3); MM OC-3 Dual-Phy ATM (Hardware 3.0; Software 11.3(5))
6. Catalyst 5500 Inge (LEC ELAN 2, MPC-B): 10/100BaseTX Supervis (Hardware 3.1; Software 4.5(2)); 10/100BaseTX Ethernet (Hardware 1.2; Software 4.5(2)); MM OC-3 Dual-Phy ATM (Hardware 2.0; Software 11.3(8)); ASP/SRP
7. Catalyst 5500 "C50-01" (LECS, LES, BUS, MPS, Router): 100BaseFX MM Supervis (Hardware 2.2; Software 4.5(3)); 10/100BaseTX Ethernet (Hardware 1.2; Software 4.5(3)); Route Switch Ext Port; Route Switch (Hardware 7.0; Software 12.0(6.5)); MM OC-3 Dual-Phy ATM (Hardware 2.0; Software 11.3(10); ASP/SRP

8. ATM NIC from ForeRunnerLE series with throughput of 155 Mbps, bus architecture PCI 2.0/2.1, 32-Bit, 33 MHz PCI Bus and Windows95/NT 3.51/4.0, NDIS 4.0 und NDIS 5.0
9. ATM NIC from Olicom Rapid Fire OC-6162: 155 Mbps over MMF (SC Connector), NDIS 3.0 NIC driver

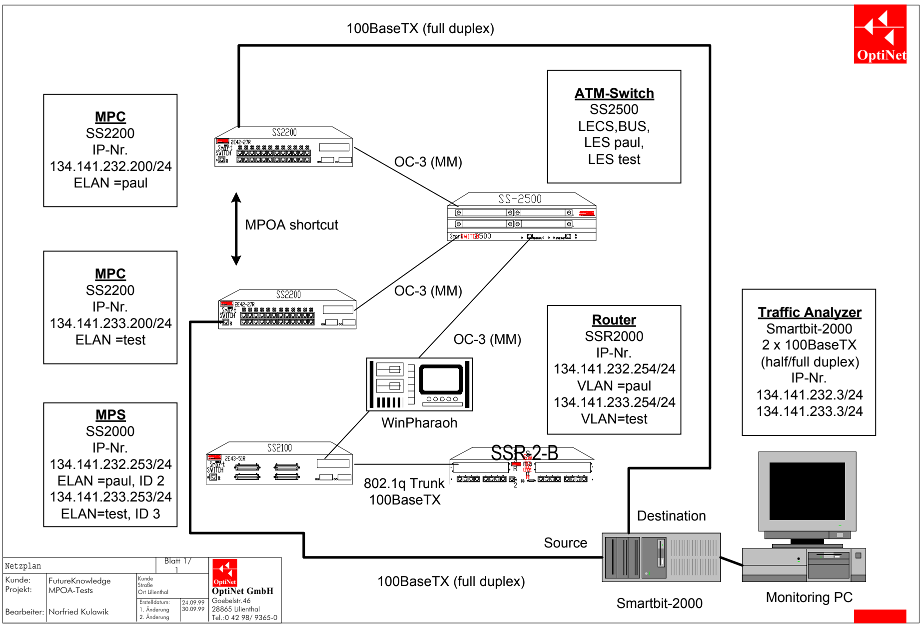

Fig. 1. MPOA scenario from Cabletron Systems (Enterasys): For the measurements with and without MPOA shortcuts, we used different shortcut frame counter for the MPOA devices. This value is responsible for the number of frames per second (fps) for the threshold, which establish a shortcut. During the tests the shortcut threshold was 10. This scenario makes it possible to establish the shortcut directly between MPC 1 and 2. On the other hand, the traffic between the SSR and ATM switch has been controlled by the Win Pharaoh. The Smartbit equipment was responsible for the traffic generation and the analysis of the test information.

As measurement equipment we used the following software and hardware from GN Nettest and NetCom:

1. Smartbits 2000 (SMB-2000): Firmware 6.220012; 2 x ML-7710 (Ethernet); 2 x AT-9155C(ATM); 1 x AT-9622(ATM); Software: SmartApplications ver. 2.22; Smart Flow 1.00.010 Beta; Library: 3.07-48; SmartSignaling 2.1(Missing authorization file, will run in demo mode); SmartWindow 6.51
2. Win Pharaoh: Hardware: LAN, WAN and ATM Line Interfaces, ISA-Bus; Prozessor: 10 MIPS RISC Prozessor; 16 MByte On-Board RAM; Software: based on Windows; works on a laptop, PC or Rack-Mount-PC; ATM Remote Software; ATM Site License; ATM Corporate License; Adapter: LAN: Fast-Ethernet, Token Ring and FDDI; WAN: RS-232, RS-422, RS-449, RS-530, V.35, X.21, V.10, V.11, Basic Rate ISDN, Primary Rate ISDN/T1 and Primary Rate ISDN/E1; ATM:

155 MBit/s OC3c/STM-1 single mode and multi mode, 155 MBit/s UTP-5, DS3/DS1, E3/E1

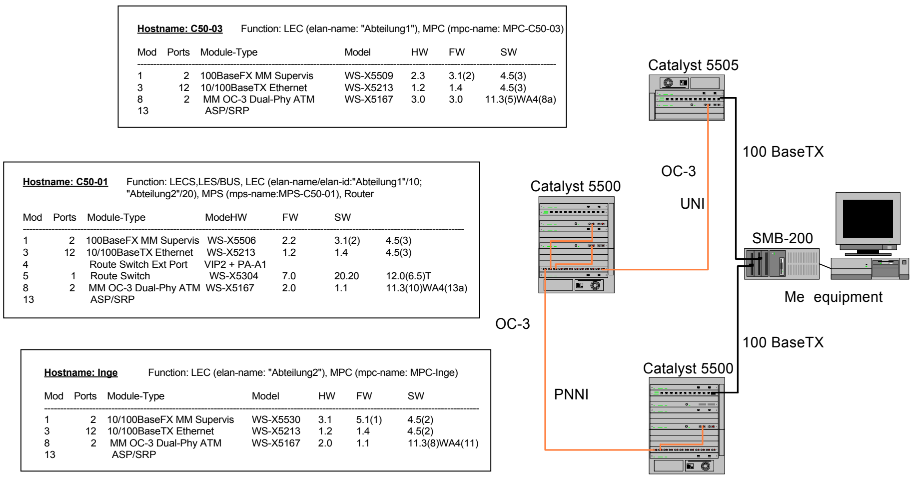

Fig. 2. MPOA scenario from Cisco Systems: The measurements with the components from Cisco Systems were realised with the Catalyst5xxx series. This are the only switches from Cisco, which support MPOA. Additionally, the router Cisco 4500, 4700, 7200, and 7500 have been integrated MPOA, too. For the tests two Catalysts worked as MPC, LEC, and Ethernet access switch. Furthermore two devices worked as MPS, LECS, LES/BUS, and router. Because of the close time frame, there is only a MPOA shortcut within the ATM switch possible. For the preparation of a MPS a RSM with VIP2 (Versatile Interface Processor) module with one ATM NIC (PA-Ax) is necessary. Two modules were devided one slot and were connected via the backplane with the Supervisor Engine and the RSM.

2.3 TCP/IP-over-ATM Scenario

The tests were carried out over a 155 Mbps connection between two separate Pentium -II/350MHz and Pentium-II/266MHz workstations with operating systems Windows98 and Windows NT 4.0. Both clients had an ATM Network Interface Card (NIC) from Fore Systems (ForeRunnerLE155) and also the RapidFire6162 ATM 155 PCI Adapter from Olicom. The central ATM switch was Cabletron's SmartSwitch2000 with a single ATM module and four OC-3 interfaces. AAL-5 was used for all measurements to encapsulated IP packets. The sender buffer varied between 16-64 kbyte, whilst the receiver buffer was always of the same value. Classical IP (CLIP) was used for the scenario without any routing. Therefore, it was a point-to-point connection between both workstations under optimal conditions.

Netperf, originally developed by Hewlett Packard, was the software used for test-purposes on both clients. It is a benchmark program that to measure the performance of networks. Measurement opportunities for burst traffic and request/response performance are the main tasks of the program. Optional extensions of the software are available. In contrast to the full Netperf version that is only available for Unix operation systems, the used test-version was especially designed for Windows end systems.

2.3.1 MPOA Measurement Results

Cabletron Systems (Enterasys) latency test without MPOA: The measurements were done by increasing of the throughput by 10 Mbps steps up to 100 Mbps. The duration of one test run was 10 sec, because of the total time of the measurements. In this diagram we used unidirectional connections established by the Smartbit.

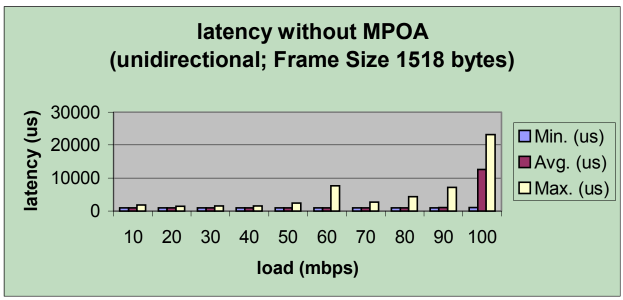

Fig. 3. The frame size differ between 64 to 1518 byte. Here you see only the 1518 byte test results, because is the significant size in the network area. Appropriate without MPOA there should not appear high latency. This is the fact, because only at 90-100% load there were higher values. In this case, the test results differ between zero and approx. 23000 ∝sec.

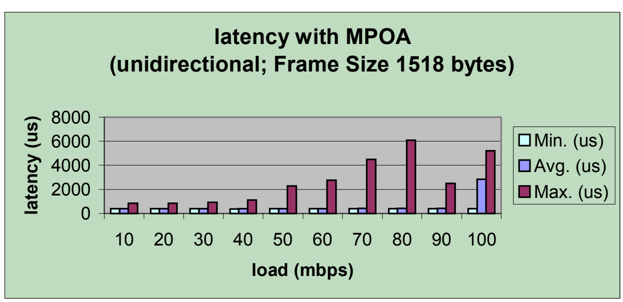

Fig. 4. Cabletron Systems (Enterasys) latency test with MPOA: The direct connection between virtual clients is different, established by a shortcut with MPOA. Here were only approx. 6000 ∝sec latency measured as maximum delay time. This was noticed at a load of 80%. Therefore, the delay has been minimised by MPOA from approx. 23000 to 6000 ∝sec.

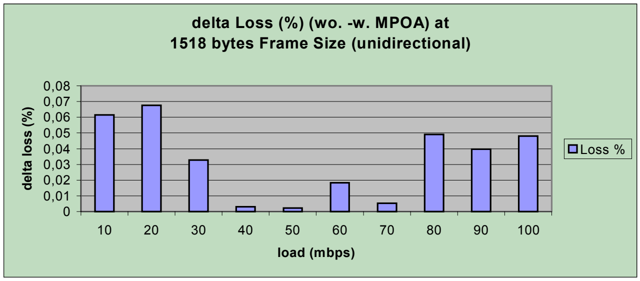

Fig. 5. Cabletron Systems (Enterasys) packet loss test with/without MPOA: If you measure the delay, you have also to keep in mind the frame losses. Here you can see the test results for Cabletron's components for MPOA (w. = with) and without MPOA (wo. = without). For packet sizes of 64 byte, there was no packet losses recorded. If there is a load of 70% there was in every case no packet losses. The value of packet losses was under 1% in every case of this measurement.

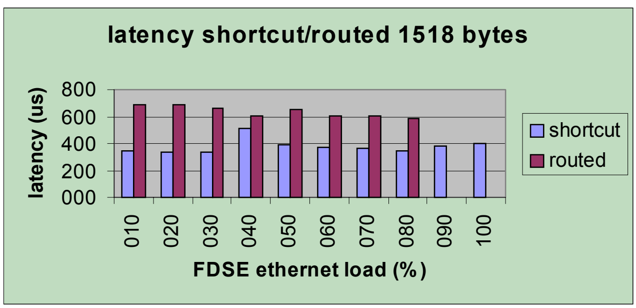

Fig. 6. Cisco Systems latency test with/without MPOA: In the first step, the latency was measured with MPOA and a load of 10-100%. Here we divided the packet sizes in 64, 128, 256, 512, 1024, 1280, and 1518 byte. So-called Under-runs (frame sizes under 64 byte) were rejected, while Giants (frame sizes over 1518 byte) were ignored by the LANE module. In the second step, the latency of the routed path was measured. Here, the router showed after a frame frequency of approx. 65000 fps, that packets were deleted. This was the reason, why no test results are there. According to Cisco Systems, this behaviour was caused by the ATM module PA-A1, which has not sufficient performance.

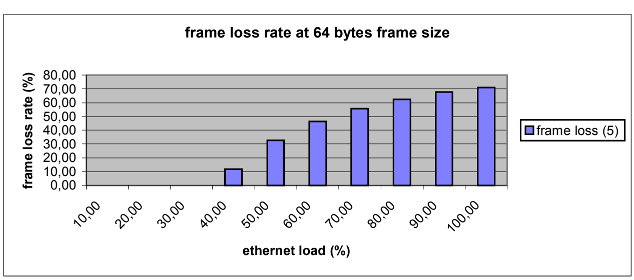

Fig. 7. Cisco Systems packet loss test with/without MPOA: In the direct comparison between with and without MPOA, you can see at 64 byte, that the routed packets need essential more time than the shortcut packets. After a load of 40% there were not measured any values, because of the frame losses.

2.4 TCP/IP-over-ATM Performance

The measurement of TCP/IP-over-ATM was carried out by using various packet sizes in order to represent the effectiveness of IP-over-ATM. The throughput alone was not really interesting, because of the different bottlenecks that were already mentioned before. The fluctuations and throughput breaks were more important. Classical IP (CLIP) has been used for this measurements point-to-point, because of the highest effectiveness. The test phase had a time duration of 60 seconds for each measurement via Netperf.

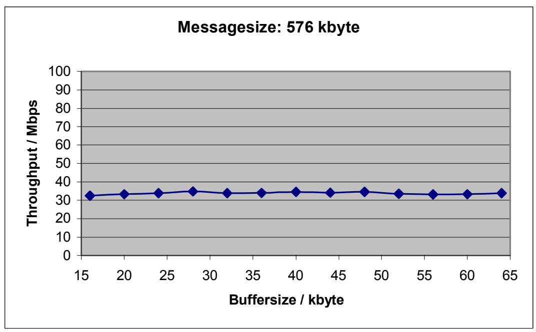

Fig. 8. IP-over-ATM, 576 byte: This figure shows the TCP/IP measurements with a packet size of 576 byte via CLIP. The buffer varies from 15-64 kbyte. The packet size of 576 byte represents a normal datagram in the Internet environment. Figure 8 demonstrates the effectiveness of the achieved throughput during a normal point-to-point session with minimal overhead. By the less datagram size the effectiveness went down to approx. 35 Mbps.

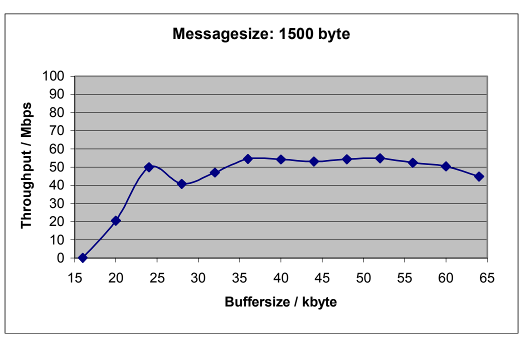

Fig. 9. IP-over-ATM, 1500 byte: Additional breaks and interruptions happened with small and big packet sizes. The MTU of 1.500 byte required a fragmentation of the packets and this has to be considered the reason for it. If the fragmentation was increased the performance went down. Figure 9 shows a differing result. In this case, the packet size was 1.500 byte which allows higher data rates then before. The results are only limited to approx. 60 Mbps. The measurements show some fluctuations, because of the different buffer sizes, which was used.

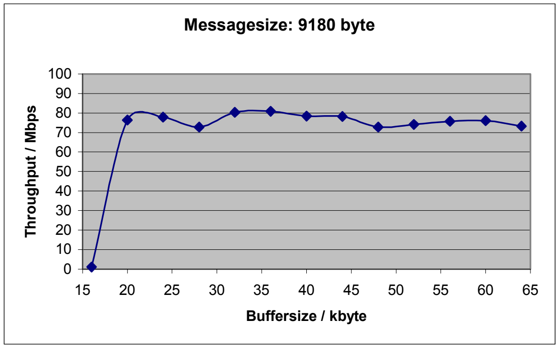

Fig. 10. Fluctuations happened, especially at a buffer size of 55 kbyte. The best results were achieved at 40 and 65 kbyte. A fundamentally improved performance was achieved with a packet size of 9.180 byte. The throughput increased up to 80 Mbps. Fluctuations happened, again (especially on less buffer sizes), like during the other measurements. The small fragmentation of the packets has to be considered the reason as the MTU size was 9.180 byte.

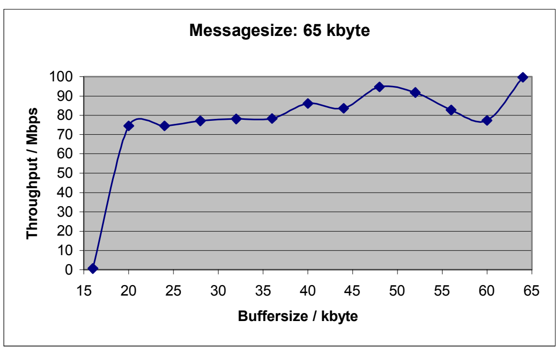

Fig. 11. IP-over-ATM, 65 kbyte: IP-over-ATM, 9180 byte: For the last measurement a packet size of 65 kbyte was chosen. It triggered several fluctuations as figure 11 shows. At first the throughput went down to approx. 5 Mbps. This performance has to be considered poor and was not efficient enough to be used in practice. Otherwise this results show the best data rates till 100 Mbps. Concluding, the big packets are useful for data networks with big buffer sizes and the maximum transport unit (MTU) is responsible for an effective data throughput.

The NIC RapidFire6162 from Olicom together with the LE155 from Fore Systems have been an efficient throughput and high Switched Virtual Circuit (SVC) performance. The LE155 has been also an active processor for direct traffic support without influence the CPU of the workstation. Both NICs need 128 Mbyte for a efficient operating. Additionally, both cards support in MPOA mode different service classes as Constant Bit Rate (CBR), Variable Bit Rate (VBR), and Unspecified Bit Rate (UBR). Available Bit Rate (ABR) was only available at the RapidFire6162. But, for the support of ABR you need this service inside the ATM switch, which was not integrated in every component yet. [3]

3. Conclusions

Original, more vendors like Olicom, IBM, 3Com, and Nortel Networks promised to take part on the tests of the WWL. But on the one hand, some manufacturers developed at the time of the test MPOA, other companies go to MPLS instead of the MPOA approach. The focus of Cisco & Co. has been changed, because the switches will be developed for frame switching and not for cell switching. [4]

Cabletron's components worked sufficient in the test environment. The shortcut functionality has been established after five packets. Therefore after 10 sec there were a shortcut established, which work with MPOA. By the measurement equipment of GN Nettest, the Win Pharaoh controlled the main connection in real-time. Netcom's Smartbits simulated and analyzed MPOA and turn over the advantages of MPOA in the LAN area. Fast layer-3-switching was possible and more effective for different packet sizes than without MPOA. The solution from Cabletron includes one more feature: the router works as layer-3-switch with IEEE802.1p and adapt the different approaches of QoS and CoS. Therefore, you can work with Gigabit-Ethernet and ATM networks.

The components of Cisco Systems established a shortcut after 10 packets. Additional, the shortcut functionality worked also efficient and sufficient. Only the blockade of the router if the system was overloaded prevented to get the test results. But the other test results were in the foreseen parameters. Therefore, the tests of the components of Cisco and the MPOA approach were successful and manageable.

TCP has been extended and further developed for better and more efficient mechanisms in high-speed networks. Yet in practice, TCP quite often does not have the optimal throughput. Several parameters, devices and mechanisms were responsible for this effect. This caused fluctuations and delays during the transmission of data traffic. Mainly TCP's own mechanisms were responsible for the small throughput such as acknowledgements, small window sizes and sequence number overloads. Additionally, the performance of the ATM boards and the end stations were important. The adapter boards from Fore Systems and Olicom showed only a maximum throughput of approx. 100 Mbps, if you use CLIP. Therefore, there are three main issues regarding latency in Internet protocols, which are able to improve the performance:

1. The default window size in many TCP/IP protocol implementations acts as a bottleneck on communications over high-latency links. On many implementations, the standard window prevents sending enough data to fill a high-latency connection.

2. TCP includes two essential congestion control mechanisms called slow start and congestion avoidance . These mean that all Internet connections (such as viewing web pages and sending e-mail) start out at lower bandwidth and then throttle up to higher speed if no congestion is encountered. The problem is that each cycle of speed increase requires a full round-trip communication between sender and receiver, and dozens of such round-trips can be necessary to reach the full potential of a link.

3. There are research efforts to look at increasing the performance of TCP by tricking the connection on the other side into believing it is communicating over a low-latency link. Unfortunately, these schemes fundamentally alter the semantics of TCP communications, introducing the possibility of data corruption. Moreover, they are incompatible with the IP security protocols (IPsec), which promise to bring an unprecedented and badly needed degree of security to the Internet

In order to implement IP-over-ATM effectively, it is necessary to assign enough time to the configuration of the participating devices and software. ATM is a new technology which considerably extents the usable bandwidth with regard to QoS parameters. Nowadays, workstations are designed for work in traditional networks such as Ethernet or Token Ring. These deadlocks must be compensated in order to use efficiently IP-over-ATM. [2, 3]

ATM is stuck in the data center, since deploying ATM to either the wiring closet or the desktop is more trouble than its worth. Some ATM vendors are trying to help this issue by building their switches to require only software upgrades for new improvements, but certain upgrades-like moving from OC-12 to OC-48 and beyond-will always require additional hardware. But you can’t count ATM out completely. It still retains the strengths its always had, namely fault tolerance and reliability. ATM's cell-based traffic and dynamic routing capabilities make it a much more resilient backbone technology than Gigabit Ethernet and considerably faster than FDDI. And, it has little if any distance limitations, which makes it a great WAN connector. Obviously then, we will continue to see ATM in the service provider arena, both in telecommunications as well as ASPs, and even large ISPs. If your corporate network especially requires these strengths, then ATM is probably in the best shape of its life as far as the enterprise is concerned. But the ATM solution will be build without MPOA and with MPLS to establish QoS mechanisms, because MPOA works but is not widespread. [4]

References

1. Detken, K.-O.: Interworking in heterogeneous environment: Multiprotocol over ATM (MPOA); European conference Interworking98; Ottawa 1998
2. Detken, K.-O.: ATM Handbook; publishing house Hüthig; Heidelberg 1999
3. Detken, K.-O.: Quality-of-Service (QoS) in heterogeneous networks: CLIP, LANE , and MPOA performance test; Networking 2000; Paris 2000
4. Rist, Oliver: ATM Still Around, But Struggling; Internet Week 10/99; USA 1999